

2/22 H2



OIPE

RAW SEQUENCE LISTING

PATENT APPLICATION: US/10/029,382

DATE: 02/28/2002

TIME: 10:30:07

Input Set : N:\Crf3\RULE60\10029382.raw.txt

Output Set: N:\CRF3\02282002\J029382.raw

1 <110> APPLICANT: DIVERSA CORPORATION
 2 CALLEN, Walter
 3 MATHUR, Eric
 4 <120> TITLE OF INVENTION: ENZYMES HAVING HIGH TEMPERATURE POLYMERASE ACTIVITY AND
 METHODS OF USE
 5 THEREOF
 6 <130> FILE REFERENCE: DIVER1350-2
 8 <140> CURRENT APPLICATION NUMBER: 10/029,382
 9 <141> CURRENT FILING DATE: 2001-12-21
 11 <150> PRIOR APPLICATION NUMBER: US/09/656,309
 12 <151> PRIOR FILING DATE: 2000-09-06
 14 <160> NUMBER OF SEQ ID NOS: 2
 15 <170> SOFTWARE: PatentIn version 3.0
 17 <210> SEQ ID NO: 1
 18 <211> LENGTH: 2412
 19 <212> TYPE: DNA
 20 <213> ORGANISM: Pyrolobus fumarius
 21 <220> FEATURE:
 22 <221> NAME/KEY: CDS
 23 <222> LOCATION: (1)..(2412)
 24 <400> SEQUENCE: 1

25	atg act gaa gtt gta ttc acg gtt tta gac tct agc tac gag gtt gtt	48
26	Met Thr Glu Val Val Phe Thr Val Leu Asp Ser Ser Tyr Glu Val Val	
27	1 5 10 15	
28	ggt aaa gag cct cag gta atc ata tgg ggt att gct gag aac ggc gag	96
29	Gly Lys Glu Pro Gln Val Ile Ile Trp Gly Ile Ala Glu Asn Gly Glu	
30	20 25 30	
31	agg gta gtc ctc att gac agg tct ttt cgc cca tac ttc tat gcg ctg	144
32	Arg Val Val Leu Ile Asp Arg Ser Phe Arg Pro Tyr Phe Tyr Ala Leu	
33	35 40 45	
34	ctt gca ccg ggc gcc gat cct aag cag gta gca caa cgt att cgt gca	192
35	Leu Ala Pro Gly Ala Asp Pro Lys Gln Val Ala Gln Arg Ile Arg Ala	
36	50 55 60	
37	ttg agt agg cca aag agc ccg att ata ggt gta gag gat gac aag agg	240
38	Leu Ser Arg Pro Lys Ser Pro Ile Ile Gly Val Glu Asp Asp Lys Arg	
39	65 70 75 80	
40	aag tac ttc ggg agg cct cgt agg gtc tta cgt att cgc acc gtg cta	288
41	Lys Tyr Phe Gly Arg Pro Arg Arg Val Leu Arg Ile Arg Thr Val Leu	
42	85 90 95	
43	ccc gag gct gtt agg gag tat cgc gaa ctc gta aag aac gtt gat ggt	336
44	Pro Glu Ala Val Arg Glu Tyr Arg Glu Leu Val Lys Asn Val Asp Gly	
45	100 105 110	
46	gtt gag gat gtt cta gag gcg gat ata cgc ttc gct atg cgc tat ctc	384
47	Val Glu Asp Val Leu Glu Ala Asp Ile Arg Phe Ala Met Arg Tyr Leu	

ENTERED

DATE: 02/28/2002

TIME: 10:30:07

Input Set : N:\Crf3\RULE60\10029382.raw.txt

Output Set: N:\CRF3\02282002\J029382.raw

48			115				120					125					
49	ata	gat	cac	gat	cta	ttt	cct	ttc	acc	tgg	tac	cgt	gta	gag	gct	gag	432
50	Ile	Asp	His	Asp	Leu	Phe	Pro	Phe	Thr	Trp	Tyr	Arg	Val	Glu	Ala	Glu	
51			130				135					140					
52	ccc	ctc	gag	aac	aag	atg	ggc	ttc	cgt	gtc	gac	aag	gta	tac	ctg	gtt	480
53	Pro	Leu	Glu	Asn	Lys	Met	Gly	Phe	Arg	Val	Asp	Lys	Val	Tyr	Leu	Val	
54			145				150					155				160	
55	aag	agc	agg	ccg	gag	cca	ctt	tat	ggg	gag	gct	ctc	gca	cca	acc	aag	528
56	Lys	Ser	Arg	Pro	Glu	Pro	Leu	Tyr	Gly	Glu	Ala	Leu	Ala	Pro	Thr	Lys	
57						165						170				175	
58	ctt	ccc	gat	ctt	agg	ata	ctc	gcg	ttc	gat	att	gaa	gtt	tat	agc	aag	576
59	Leu	Pro	Asp	Leu	Arg	Ile	Leu	Ala	Phe	Asp	Ile	Glu	Val	Tyr	Ser	Lys	
60						180						185				190	
61	caa	ggg	tcg	ccg	cgt	cca	gag	cgc	gat	cct	gta	ata	gtg	ata	gct	gtg	624
62	Gln	Gly	Ser	Pro	Arg	Pro	Glu	Arg	Asp	Pro	Val	Ile	Val	Ile	Ala	Val	
63						195							205				
64	aag	act	gac	gat	ggc	gat	gag	gtg	cta	ttc	att	gca	gag	ggc	aaa	gac	672
65	Lys	Thr	Asp	Asp	Gly	Asp	Glu	Val	Leu	Phe	Ile	Ala	Glu	Gly	Lys	Asp	
66							215						220				
67	gat	cga	aaa	ccg	ata	cgc	gag	ttt	gta	gag	tac	gtg	aag	agg	tat	gac	720
68	Asp	Arg	Lys	Pro	Ile	Arg	Glu	Phe	Val	Glu	Tyr	Val	Lys	Arg	Tyr	Asp	
69							230						235			240	
70	ccc	gac	ata	ata	gtc	ggg	tat	aac	aac	aat	cat	ttc	gat	tgg	cct	tat	768
71	Pro	Asp	Ile	Ile	Val	Gly	Tyr	Asn	Asn	Asn	His	Phe	Asp	Trp	Pro	Tyr	
72						245							250			255	
73	ctt	ttg	agg	cgc	gcc	cgc	atc	cta	ggc	ata	aag	ctt	gat	gtg	act	aga	816
74	Leu	Leu	Arg	Arg	Ala	Arg	Ile	Leu	Gly	Ile	Lys	Leu	Asp	Val	Thr	Arg	
75						260							265			270	
76	aga	gtt	ggc	gcc	gag	ccc	acc	act	agc	gta	cat	ggg	cac	gtc	tct	gtc	864
77	Arg	Val	Gly	Ala	Glu	Pro	Thr	Ser	Val	His	Gly		His	Val	Ser	Val	
78													285				
79	cct	ggc	agg	ctt	aac	gta	gat	ctg	tac	gac	tat	gcc	gaa	gag	atg	cca	912
80	Pro	Gly	Arg	Leu	Asn	Val	Asp	Leu	Tyr	Asp	Tyr	Ala	Glu	Glu	Met	Pro	
81							295						300				
82	gag	atc	aag	ata	aag	agt	ctc	gag	gag	gtc	gca	gag	tat	cta	ggc	gtg	960
83	Glu	Ile	Lys	Ile	Lys	Ser	Leu	Glu	Glu	Val	Ala	Glu	Tyr	Leu	Gly	Val	
84							310						315			320	
85	atg	aag	aag	agt	gaa	cgc	gtt	atc	atc	aat							

RAW SEQUENCE LISTING

DATE: 02/28/2002

PATENT APPLICATION: US/10/029,382

TIME: 10:30:07

Input Set : N:\Crf3\RULE60\10029382.raw.txt

Output Set: N:\CRF3\02282002\J029382.raw

97	gcg atg agt gtt ggc ttt cga ctt gaa tgg tac ctg ata cgc gcg gcg	1200
98	Ala Met Ser Val Gly Phe Arg Leu Glu Trp Tyr Leu Ile Arg Ala Ala	
99	385 390 395 400	
100	ttt aag atg aaa gag ctt gtg ccg aac cgc gtt gag cgc cca gaa gag	1248
101	Phe Lys Met Lys Glu Leu Val Pro Asn Arg Val Glu Arg Pro Glu Glu	
102	405 410 415	
103	act tac cgt ggc gct ata gtt ctt gag ccg ttg aga ggc gtg cac gag	1296
104	Thr Tyr Arg Gly Ala Ile Val Leu Glu Pro Leu Arg Gly Val His Glu	
105	420 425 430	
106	aat ata gcc gta ctc gac ttt agc tcg atg tac cca aac atc atg ata	1344
107	Asn Ile Ala Val Leu Asp Phe Ser Ser Met Tyr Pro Asn Ile Met Ile	
108	435 440 445	
109	aag tac aat gtt ggt cct gac acg ctt gtg agg cct ggt gaa aag tgt	1392
110	Lys Tyr Asn Val Gly Pro Asp Thr Leu Val Arg Pro Gly Glu Lys Cys	
111	450 455 460	
112	ggc gag tgt ggt tgc tgg gag gcc ccg gag gtc aag cac agg ttc cgt	1440
113	Gly Glu Cys Gly Cys Trp Glu Ala Pro Glu Val Lys His Arg Phe Arg	
114	465 470 475 480	
115	agg tgt ccg ccc ggc ttc ttc aag aca gtt ctt gag agg ctg tta gag	1488
116	Arg Cys Pro Pro Gly Phe Phe Lys Thr Val Leu Glu Arg Leu Leu Glu	
117	485 490 495	
118	ctt cgt aag cgt gtg cgt gct gaa atg aag aag tat cct ccg gat agc	1536
119	Leu Arg Lys Arg Val Arg Ala Glu Met Lys Lys Tyr Pro Pro Asp Ser	
120	500 505 510	
121	cca gaa tat cga ctg ttg gat gaa agg cag aag gcg ttg aag gtt ctt	1584
122	Pro Glu Tyr Arg Leu Leu Asp Glu Arg Gln Lys Ala Leu Lys Val Leu	
123	515 520 525	
124	gca aac gct agt tac ggc tac atg ggt tgg agc ggc gct agg tgg tat	1632
125	Ala Asn Ala Ser Tyr Gly Tyr Met Gly Trp Ser Gly Ala Arg Trp Tyr	
126	530 535 540	
127	tgc agg gag tgc gca aag gct gtc acg gct tgg ggt agg cac ctc ata	1680
128	Cys Arg Glu Cys Ala Lys Ala Val Thr Ala Trp Gly Arg His Leu Ile	
129	545 550 555 560	
130	cgc acc gcc atc aac ata gct cgt aaa cta ggc ctc aag gtg atc tac	1728
131	Arg Thr Ala Ile Asn Ile Ala Arg Lys Leu Gly Leu Lys Val Ile Tyr	
132	565 570 575	
133	ggt gac aca gat tcg ctc ttc gtg acc tat gat ccg gag aag gtg gaa	1776
134	Gly Asp Thr Asp Ser Leu Phe Val Thr Tyr Asp Pro Glu Lys Val Glu	
135	580 585 590	
136	aat ttc atc aaa att ata aag gag gag ctg ggg ttc gaa atc aag cta	1824
137	Asn Phe Ile Lys Ile Ile Lys Glu Glu Leu Gly Phe Glu Ile Lys Leu	
138	595 600 605	
139	gag aag gtg tac aaa cgc tta ttc ttt aca gag gct aag aag agg tac	1872
140	Glu Lys Val Tyr Lys Arg Leu Phe Phe Thr Glu Ala Lys Lys Arg Tyr	
141	610 615 620	
142	gct ggc ctt ctc gag gac gga cgt ata gat att gtc ggt ttc gag gct	1920
143	Ala Gly Leu Leu Glu Asp Gly Arg Ile Asp Ile Val Gly Phe Glu Ala	
144	625 630 635 640	
145	gta cgt ggc gat tgg tgt gaa ctc gcc aag gag gtt cag act aag gtt	1968

RAW SEQUENCE LISTING

DATE: 02/28/2002

PATENT APPLICATION: US/10/029,382

TIME: 10:30:07

Input Set : N:\Crf3\RULE60\10029382.raw.txt

Output Set: N:\CRF3\02282002\J029382.raw

146	Val Arg Gly Asp Trp Cys Glu Leu Ala Lys Glu Val Gln Thr Lys Val	
147	645 650 655	
148	gtc gaa ata gta ttg aag acg agt gag gtg aac aag gct gta gag tac	2016
149	Val Glu Ile Val Leu Lys Thr Ser Glu Val Asn Lys Ala Val Glu Tyr	
150	660 665 670	
151	gtc agg aag att gtg aaa gag ttg gag gag ggc aag gtt ccc ata gag	2064
152	Val Arg Lys Ile Val Lys Glu Leu Glu Glu Gly Lys Val Pro Ile Glu	
153	675 680 685	
154	aag ctt gta atc tgg aag acc ctt agt aag cgt ctt gag gag tac aca	2112
155	Lys Leu Val Ile Trp Lys Thr Leu Ser Lys Arg Leu Glu Glu Tyr Thr	
156	690 695 700	
157	acg gag gca cca cac gtc gtt gca gcg aag agg atg ctg tca gca ggc	2160
158	Thr Glu Ala Pro His Val Val Ala Ala Lys Arg Met Leu Ser Ala Gly	
159	705 710 715 720	
160	tac cgg gta agc cca ggc gac aag ata ggg tat gta ata gtg aag ggt	2208
161	Tyr Arg Val Ser Pro Gly Asp Lys Ile Gly Tyr Val Ile Val Lys Gly	
162	725 730 735	
163	ggg ggc cgt atc agt caa aga gca tgg cca tac ttc atg gtc aag gat	2256
164	Gly Gly Arg Ile Ser Gln Arg Ala Trp Pro Tyr Phe Met Val Lys Asp	
165	740 745 750	
166	cct agc cag ata gac gtg acc tac tat gtt gac cac caa atc atc ccg	2304
167	Pro Ser Gln Ile Asp Val Thr Tyr Tyr Val Asp His Gln Ile Ile Pro	
168	755 760 765	
169	gct gca ttg aga ata ctg ggc tac ttt ggc atc acc gag aag aag ctg	2352
170	Ala Ala Leu Arg Ile Leu Gly Tyr Phe Gly Ile Thr Glu Lys Lys Leu	
171	770 775 780	
172	aaa gca agt gca act ggg cag aag act ctc ttc gac ttt cta gcc aag	2400
173	Lys Ala Ser Ala Thr Gly Gln Lys Thr Leu Phe Asp Phe Leu Ala Lys	
174	785 790 795 800	
175	aag agc aag taa	2412
176	Lys Ser Lys	
178	<210> SEQ ID NO: 2	
179	<211> LENGTH: 803	
180	<212> TYPE: PRT	
181	<213> ORGANISM: Pyrolobus fumarius	
182	<400> SEQUENCE: 2	
183	Met Thr Glu Val Val Phe Thr Val Leu Asp Ser Ser Tyr Glu Val Val	
184	1 5 10 15	
185	Gly Lys Glu Pro Gln Val Ile Ile Trp Gly Ile Ala Glu Asn Gly Glu	
186	20 25 30	
187	Arg Val Val Leu Ile Asp Arg Ser Phe Arg Pro Tyr Phe Tyr Ala Leu	
188	35 40 45	
189	Leu Ala Pro Gly Ala Asp Pro Lys Gln Val Ala Gln Arg Ile Arg Ala	
190	50 55 60	
191	Leu Ser Arg Pro Lys Ser Pro Ile Ile Gly Val Glu Asp Asp Lys Arg	
192	65 70 75 80	
193	Lys Tyr Phe Gly Arg Pro Arg Arg Val Leu Arg Ile Arg Thr Val Leu	
194	85 90 95	
195	Pro Glu Ala Val Arg Glu Tyr Arg Glu Leu Val Lys Asn Val Asp Gly	

RAW SEQUENCE LISTING

PATENT APPLICATION: US/10/029,382

DATE: 02/28/2002
TIME: 10:30:07Input Set : N:\Crf3\RULE60\10029382.raw.txt
Output Set: N:\CRF3\02282002\J029382.raw

196 Val Glu Asp Val Leu Glu Ala Asp Ile Arg Phe Ala Met Arg Tyr Leu 110
197 115
198 Ile Asp His Asp Leu Phe Pro Phe Thr Trp Tyr Arg Val Glu Ala Glu 125
199 130
200 Pro Leu Glu Asn Lys Met Gly Phe Arg Val Asp Lys Val Tyr Leu Val 140
201 145
202 Lys Ser Arg Pro Glu Pro Leu Tyr Gly Glu Ala Leu Ala Pro Thr Lys 155
203 165
204 Leu Pro Asp Leu Arg Ile Leu Ala Phe Asp Ile Glu Val Tyr Ser Lys 170
205 180
206 Gln Gly Ser Pro Arg Pro Glu Arg Asp Pro Val Ile Val Ile Ala Val 185
207 195
208 Lys Thr Asp Asp Gly Asp Glu Val Leu Phe Ile Ala Glu Gly Lys Asp 205
209 210
210 Asp Arg Lys Pro Ile Arg Glu Phe Val Glu Tyr Val Lys Arg Tyr Asp 220
211 225
212 Pro Asp Ile Ile Val Gly Tyr Asn Asn Asn His Phe Asp Trp Pro Tyr 235
213 245
214 Leu Leu Arg Arg Ala Arg Ile Leu Gly Ile Lys Leu Asp Val Thr Arg 250
215 260
216 Arg Val Gly Ala Glu Pro Thr Thr Ser Val His Gly His Val Ser Val 265
217 275
218 Pro Gly Arg Leu Asn Val Asp Leu Tyr Asp Tyr Ala Glu Glu Met Pro 280
219 290
220 Glu Ile Lys Ile Lys Ser Leu Glu Glu Val Ala Glu Tyr Leu Gly Val 295
221 305
222 Met Lys Lys Ser Glu Arg Val Ile Ile Asn Trp Trp Glu Ile Pro Asp 310
223 325
224 Tyr Trp Asp Asp Pro Lys Lys Arg Pro Leu Leu Gln Tyr Ala Arg 330
225 340
226 Asp Asp Val Arg Ala Thr Tyr Gly Leu Ala Glu Lys Ile Leu Pro Phe 345
227 355
228 Ala Ile Gln Leu Ser Tyr Val Thr Gly Leu Pro Leu Asp Gln Val Gly 360
229 370
230 Ala Met Ser Val Gly Phe Arg Leu Glu Trp Tyr Leu Ile Arg Ala Ala 375
231 385
232 Phe Lys Met Lys Glu Leu Val Pro Asn Arg Val Glu Arg Pro Glu Glu 390
233 405
234 Thr Tyr Arg Gly Ala Ile Val Leu Glu Pro Leu Arg Gly Val His Glu 410
235 420
236 Asn Ile Ala Val Leu Asp Phe Ser Ser Met Tyr Pro Asn Ile Met Ile 425
237 435
238 Lys Tyr Asn Val Gly Pro Asp Thr Leu Val Arg Pro Gly Glu Lys Cys 440
239 450
240 Gly Glu Cys Gly Cys Trp Glu Ala Pro Glu Val Lys His Arg Phe Arg 455
241 465
242 Arg Cys Pro Pro Gly Phe Phe Lys Thr Val Leu Glu Arg Leu Leu Glu 470
243 485
244 490

VERIFICATION SUMMARY

DATE: 02/28/2002

PATENT APPLICATION: US/10/029,382

TIME: 10:30:08

Input Set : N:\Crf3\RULE60\10029382.raw.txt

Output Set: N:\CRF3\02282002\J029382.raw